Optimal Pointers for Joint Measurement of σ_x and σ_z via Homodyne Detection

Bas Janssens* and Luc Bouten†

*Department of Mathematics, Radboud University Nijmegen,
Toernooiveld 1, 6525 ED, Nijmegen, The Netherlands

†Physical Measurement and Control 266-33, California Institute of Technology,
1200 E. California Blvd., Pasadena, CA 91125, USA

Abstract

We study a model of a qubit in interaction with the electromagnetic field. By means of homodyne detection, the field-quadrature $A_t + A_t^*$ is observed continuously in time. Due to the interaction, information about the initial state of the qubit is transferred into the field, thus influencing the homodyne measurement results. We construct random variables (pointers) on the probability space of homodyne measurement outcomes having distributions close to the initial distributions of σ_x and σ_z. Using variational calculus, we find the pointers that are optimal. These optimal pointers are very close to hitting the bound imposed by Heisenberg's uncertainty relation on joint measurement of two non-commuting observables. We close the paper by giving the probability densities of the pointers.

1 Introduction

The implementation of quantum filtering and control [5] in recent experiments [2], [13] has brought new interest to the field of continuous time measurement of quantum systems [11], [12], [5], [7], [10], [25], [9]. In particular, homodyne detection has played a considerable role in this development [10]. In this paper, we aim to gain insight into the transfer of information about the initial state of a qubit from this qubit, a two-level atom, to the homodyne photocurrent, which is observed in actual experiments. Our goal is to perform a joint measurement of two non-commuting observables in the initial system. In order to achieve this, we construct random variables (pointers) on the space of possible homodyne measurement results, having distributions close (in a sense to be defined) to the distributions of these observables in the initial state.

The problem of joint measurement of non-commuting observables has been studied by several authors before, see [22], [12], [16] and the references therein. As a measure for the quality of an unbiased measurement, we use the difference between the variance of the pointer in the final state and the variance of the observable in the initial state, evaluated in the worst case initial state [18]. In other words, the quality of measurement is given in terms of the worst case added variance. These worst case added variances for two pointers, corresponding to two non-commuting observables of the initial system, satisfy a Heisenberg-like relation that bounds how well their joint measurement can be performed [18].

The paper concentrates on the example of a qubit coupled to the quantized electromagnetic field. We study this system in the weak coupling limit [15], i.e. the interaction between qubit and field is governed by a quantum stochastic differential equation in the sense of Hudson and Parthasarathy [17]. In the electromagnetic field we perform a homodyne detection experiment. Its integrated photocurrent is the measurement result for measurement of the field-quadrature $A_t + A_t^*$ continuously in time. Using the characteristic functions introduced by Barchielli and Lupieri [4], we find the probability density for these measurement results. In this density the x- and z-component of the Bloch vector of the initial state appear, indicating that homodyne detection is in fact a joint measurement of σ_x and σ_z in the initial state.

The goal of the paper is to construct random variables (pointers) on the probability space of homodyne measurement results having distributions as close as possible to those of the observables σ_x and σ_z in the initial state of the qubit. 'As close as possible' is taken to mean that the pointer must give an unbiased estimate of the observable, with its worst case added variance as low as possible. Using an argument due to Wiseman [24], we first show that optimal random variables will only depend on the endpoint of a weighted path of the integrated photocurrent. Allowed to restrict our attention to this smaller class of pointers, we are able to use standard variational calculus to obtain the optimal random variables. They do not achieve the bound imposed by the Heisenberg-like relation for the worst case added variances [18], but will be off by less than 5.6%.

The remainder of the paper is organized as follows. In Section 2 we introduce the model of the qubit coupled to the field in the weak coupling limit. Section 3 introduces the quality of a measurement in terms of the worst case added variance. This section also contains the Heisenberg-like relation for joint measurement. In Section 4 we calculate the characteristic function of Barchielli and Lupieri for the homodyne detection experiment. Section 5 deals with the variational calculus to find the optimal pointers. In Section 6 we calculate the densities of the optimal pointers and then capture our main results graphically. In the last section we discuss our results.

2 The model

We consider a two-level atom, i.e. a qubit, in interaction with the quantized electromagnetic field. The qubit is described by $\mathbb{C}^2$ and the electromagnetic field by the *symmetric Fock space* $\mathcal{F}$ over the Hilbert space of quadratically integrable functions $L^2(\mathbb{R})$ (space of one-photon wave functions), i.e.

$$\mathcal{F} := \mathbb{C} \oplus \bigoplus_{k=1}^{\infty} L^2(\mathbb{R})^{\otimes_s k}.$$

With the Fock space $\mathcal{F}$ we can describe superpositions of field-states with different numbers of photons. The joint system of qubit and field together is described by the Hilbert space $\mathbb{C}^2 \otimes \mathcal{F}$.

The interaction between the qubit and the electromagnetic field is studied in the weak coupling limit [14], [15], [1]. This means that in the interaction picture the unitary dynamics of the qubit and the field together is given by a quantum stochastic differential equation (QSDE) in the sense of Hudson and Parthasarathy [17]

$$dU_t = \Big\{\sigma_- dA_t^* - \sigma_+ dA_t - \tfrac{1}{2}\sigma_+\sigma_- dt\Big\}U_t, \quad \text{with} \quad \sigma_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \ \sigma_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \ U_0 = I. \quad (1)$$

The operators σ_- and σ_+ are the annihilator and creator on the two-level system. The field annihilation and creation processes are denoted A_t and A_t^*, respectively. Keep in mind that the evolution U_t acts nontrivially on the combined system $\mathbb{C}^2 \otimes \mathcal{F}$, whereas $\sigma_\pm$ and A_t are understood to designate the single system-operators $\sigma_\pm \otimes I$ and $I \otimes A_t$. Throughout the paper we will remain in the interaction picture. Equation (1) should be understood as a shorthand for the integral equation

$$U_t = I + \int_0^t \sigma_- U_\tau dA_\tau^* - \int_0^t \sigma_+ U_\tau dA_\tau - \tfrac{1}{2}\int_0^t \sigma_+\sigma_- U_\tau d\tau,$$

where the integrals on the right-hand side are stochastic integrals in the sense of Hudson and Parthasarathy [17]. The value of these integrals does not lie in their actual definition (on which we will not comment further), but in the Itô rule satisfied by them, allowing for easy calculations.

Theorem 2.1: ***(Quantum Itô rule [17], [20])*** *Let X_t and Y_t be stochastic integrals of the form*

$$dX_t = C_t dA_t + D_t dA_t^* + E_t dt$$
$$dY_t = F_t dA_t + G_t dA_t^* + H_t dt$$

for some stochastically integrable processes C_t, D_t, E_t, F_t, G_t and H_t (see [17], [20] for definitions). Then the process X_tY_t satisfies the relation

$$d(X_tY_t) = X_t dY_t + (dX_t)Y_t + dX_t dY_t,$$

where $dX_t dY_t$ should be evaluated according to the quantum Itô table:

	dA_t	dA_t^*	dt
dA_t	0	dt	0
dA_t^*	0	0	0
dt	0	0	0

i.e. $dX_t dY_t = C_t G_t dt$.

As a corollary we have that, for any $f \in C^2(\mathbb{R})$, the process $f(X_t)$ satisfies $d(f(X_t)) = f'(X_t)dX_t + \frac{1}{2}f''(X_t)(dX_t)^2$, where $(dX_t)^2$ should be evaluated according to the quantum Itô table.

First a matter of notation. The quantum Itô rule will be used for calculating differentials of products of stochastic integrals. Let $\{Z_i\}_{i=1,\ldots,p}$ be stochastic integrals. Then we write

$$d(Z_1 Z_2 \ldots Z_p) = \sum_{\substack{\nu \subset \{1,\ldots,p\} \\ \nu \neq \emptyset}} [\nu]$$

where the sum runs over all *non-empty* subsets of $\{1,\ldots,p\}$. For any $\nu = \{i_1,\ldots,i_k\}$, the term $[\nu]$ is the contribution to $d(Z_1 Z_2 \ldots Z_p)$ coming from differentiating only the terms with indices in the set $\{i_1,\ldots,i_k\}$ and preserving the order of the factors in the product. The differential $d(Z_1 Z_2 Z_3)$, for example, contains terms of type [1], [2], [3], [12], [13], [23] and [123]. We have $[2] = Z_1(dZ_2)Z_3$, $[13] = (dZ_1)Z_2(dZ_3)$, $[123] = (dZ_1)(dZ_2)(dZ_3)$, etc.

Let us return to equation (1). In order to illustrate how the quantum Itô rule will be used, we calculate the time evolution on the qubit explicitly. The algebra of qubit-observables is the algebra of 2×2-matrices, denoted $M_2(\mathbb{C})$. The algebra of observables in the field is given by $B(\mathcal{F})$, the bounded operators on $\mathcal{F}$. If $\mathrm{id} : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ is the identity map and $\phi : B(\mathcal{F}) \to \mathbb{C}$ is the expectation with respect to the vacuum state $\Phi := 1 \oplus 0 \oplus 0 \oplus \ldots \in \mathcal{F}$ (i.e. $\phi(Y) := \langle \Phi, Y\Phi \rangle$), then time evolution on the qubit $T_t : M_2(\mathbb{C}) \to M_2(\mathbb{C})$ is given by $T_t(X) := \mathrm{id} \otimes \phi(U_t^* X \otimes I U_t)$. On the combined system, the full time evolution $j_t : M_2(\mathbb{C}) \otimes B(\mathcal{F}) \to M_2(\mathbb{C}) \otimes B(\mathcal{F})$ is given by $j_t(W) := U_t^* W U_t$. In a diagram this reads

$$\begin{tikzcd} M_2 \arrow[r, "T_t"] \arrow[d, "\mathrm{id}\otimes I"'] & M_2 \\ M_2 \otimes B(\mathcal{F}) \arrow[r, "j_t"] & M_2 \otimes B(\mathcal{F}). \arrow[u, "\mathrm{id}\otimes\phi"'] \end{tikzcd} \tag{2}$$

In the Schrödinger picture the arrows would be reversed. A qubit-state ρ would be extended with the vacuum to $\rho \otimes \phi$, then time evolved with U_t, and in the last step the partial trace over the field would be taken, resulting in the state $\rho \circ T_t$.

Using the Itô rule we can derive a (matrix-valued) differential equation for $T_t(X)$, i.e.

$$\begin{aligned} dT_t(X) &= \mathrm{id} \otimes \phi\big(d(U_t^* X \otimes I U_t)\big) \\ &= \mathrm{id} \otimes \phi\big((dU_t^*) X \otimes I U_t + U_t^* X \otimes I (dU_t) + (dU_t^*) X \otimes I (dU_t)\big) \\ &= \mathrm{id} \otimes \phi\big(U_t^* L(X) \otimes I U_t\big) dt \\ &= T_t\big(L(X)\big) dt, \end{aligned} \tag{3}$$

where L is the Lindblad generator

$$L(X) := -\tfrac{1}{2}(\sigma_+\sigma_- X + X\sigma_+\sigma_-) + \sigma_+ X \sigma_-.$$

In the derivation (3) we used the QSDE for U_t^* which easily follows from (1)

$$dU_t^* = U_t^*\Big\{\sigma_+ dA_t - \sigma_- dA_t^* - \tfrac{1}{2}\sigma_+\sigma_- dt\Big\}, \qquad U_0^* = I.$$

Furthermore, we used that stochastic integrals with respect to dA_t and dA_t^* vanish with respect to the vacuum expectation, leaving us only with the dt terms. The differential equation (3) with initial condition $T_0(X) = X$ is solved by $T_t(X) = \exp(tL)(X)$, which is exactly the time evolution of a two-level system spontaneously decaying to the ground state, as it should be. Although the arguments above are completely standard (cf. [17]), they do illustrate nicely and briefly some of the techniques used also in following sections.

3 Quality of information transfer

Now suppose we do a homodyne detection experiment, enabling us to measure the observables $A_t + A_t^*$ in the field continuously in time [3]. If initially the qubit is in state ρ, then at time t the qubit and field together are in a state ρ^t on $M_2(\mathbb{C}) \otimes B(\mathcal{F})$ given by $\rho^t(W) := \rho \otimes \phi(U_t^* W U_t) = \rho\big(\mathrm{id} \otimes \phi(U_t^* W U_t)\big)$. Since

$$\begin{aligned} d\Big(\mathrm{id} \otimes \phi(U_t^* I \otimes (A_t + A_t^*) U_t)\Big) &= \mathrm{id} \otimes \phi\Big(d(U_t^* I \otimes (A_t + A_t^*) U_t)\Big) \\ &= \mathrm{id} \otimes \phi\Big([1] + [2] + [3] + [12] + [13] + [23] + [123]\Big) \\ &= \mathrm{id} \otimes \phi\Big(U_t^* (\sigma_- + \sigma_+) \otimes I U_t\Big) dt \\ &= \exp(tL)(\sigma_- + \sigma_+) dt = e^{-\frac{t}{2}} \sigma_x dt, \end{aligned}$$

we have that regardless the initial state ρ of the qubit, the expectation of $(A_t + A_t^*)$ in the final state ρ^t will equal the expectation of $(2 - 2e^{-\frac{t}{2}})\sigma_x$ in the initial state ρ.

3.1 Defining the quality of information transfer

The process at hand is thus a transfer of information about σ_x to a 'pointer' $A_t + A_t^*$, which can be read off by means of homodyne detection. This motivates the following definition.

Definition 3.1: (Unbiased Measurement [18]) Let X be an observable of the qubit, i.e. a self-adjoint element of $M_2(\mathbb{C})$, and let Y be an observable of the field, i.e. a self-adjoint operator in (or affiliated to) $B(\mathcal{F})$. An *unbiased measurement* M of X with *pointer* Y is by definition a completely positive map $M : \ B(\mathcal{F}) \to M_2(\mathbb{C})$ such that $M(Y) = X$.

Needless to say, for each fixed t the map $M : \ B(\mathcal{F}) \to M_2(\mathbb{C})$ given by $M(B) := \mathrm{id} \otimes \phi(U_t^* I \otimes B U_t)$ is a measurement of σ_x with pointer $Y = (2 - 2e^{-\frac{t}{2}})^{-1}(A_t + A_t^*)$. This means that, after the measurement procedure of coupling to the field in the vacuum state and allowing for interaction with the qubit for t time units, the distribution of the measurement results of the pointer Y has inherited the expectation of σ_x, regardless of the initial state ρ. However, we are more ambitious and would like its distribution as a whole to resemble that of σ_x. This motivates the following definition.

Definition 3.2: (Quality [18]) Let $M : \ B(\mathcal{F}) \to M_2(\mathbb{C})$ be an unbiased measurement of X with pointer Y. Then its *quality* σ is defined by

$$\sigma^2 := \sup\Big\{\mathrm{Var}_{\rho \circ M}(Y) - \mathrm{Var}_\rho(X) \Big| \ \rho \in \mathcal{S}(M_2)\Big\},$$

where $\mathcal{S}(M_2)$ denotes the state space of $M_2(\mathbb{C})$ (i.e. all positive normalized linear functionals on $M_2(\mathbb{C})$).

This means that σ^2 is the variance added to the initial distribution of X by the measurement

procedure M for the worst case initial state ρ. A small calculation shows that

$$\mathrm{Var}_{\rho\circ M}(Y) - \mathrm{Var}_{\rho}(X) = \rho\big(M(Y^2) - M(Y)^2\big),$$

which implies that $\sigma^2 = \|M(Y^2) - M(Y)^2\|$, where $X \mapsto \|X\|$ denotes the operator norm on $M_2(\mathbb{C})$. In particular this shows that σ^2 is positive, as one might expect. It follows from [18] that σ equals zero if and only if the measurement procedure M exactly carries over the distribution of X to Y. In short, σ is a suitable measure for how well M transfers information about X to the pointer Y.

3.2 Calculating the quality of information transfer

Let us return to the example at hand, i.e. $M(B) = \mathrm{id}\otimes\phi(U_t^* I \otimes BU_t)$, with field-observable $Y = (2 - 2e^{-\frac{t}{2}})^{-1}(A_t + A_t^*)$ as a pointer for σ_x. Let us calculate its quality, which amounts to evaluating $M(Y^2) = (2-2e^{-\frac{t}{2}})^{-2}M\big((A_t + A_t^*)^2\big)$. To this aim, we will first introduce some ideas which will be of use to us in later calculations as well.

Definition 3.3: Let f and h be real valued functions, h twice differentiable. Let Y_t be given by $dY_t = f(t)(dA_t + dA_t^*)$, $Y_0 = 0$. For $X \in M_2(\mathbb{C})$ we define

$$F_h(X,t) := \mathrm{id}\otimes\phi\big(U_t^* X \otimes h(Y_t)U_t\big).$$

When no confusion can arise we shall shorten $F_h(X,t)$ to $F_h(X)$.

The homodyne detection experiment has given us a measurement result (the integrated photocurrent) which is just the path of measurement results for $A_t + A_t^*$ continuously in time. Given this result, we post-process it by weighting the increments of the path with the function $f(t)$ and letting $h(y)$ act on the result. The following lemma will considerably shorten calculations.

Lemma 3.4:

$$\frac{dF_h(X)}{dt} = F_h\big(L(X)\big) + f(t)F_{h'}(\sigma_+ X + X\sigma_-) + \tfrac{1}{2}f(t)^2 F_{h''}(X)$$

Proof. Using the notation below Theorem 2.1 with $Z_1 = U_t^*$, $Z_2 = I\otimes h(Y_t)$ and $Z_3 = U_t$, we find

$$dF_h(X) = \mathrm{id}\otimes\phi\big([1]+[2]+[3]+[12]+[13]+[23]+[123]\big).$$

Again we will use that the vacuum expectation kills all dA_t and dA_t^* terms. Using Theorem 2.1 we see that after the vacuum expectation the terms [1], [3] and [13] make up $F_h\big(L(X)\big)dt$. Since third powers of increments are 0 we again have $[123] = 0$. From

$$dh(Y_t) = h'(Y_t)f(t)(dA_t + dA_t^*) + \tfrac{1}{2}h''(Y_t)f(t)^2dt,$$

we find that, after taking vacuum expectations, the terms [12] and [23] make up the second term $f(t)F_{h'}(\sigma_+X + X\sigma_-)dt$ and [2] provides the last term $\frac{1}{2}f(t)^2F_{h''}(X)dt$. □

We are now well-equipped to calculate $M\big((A_t + A_t^*)^2\big)$. Choose $f(t) = 1$ and $h(x) = x^2$. (The maps $x \mapsto x^n$ will be denoted $\mathfrak{x}^n$ hereafter.) Then $M\big((A_t + A_t^*)^2\big) = F_{\mathfrak{x}^2}(I)$ and by Lemma 3.4

$$\frac{dF_{\mathfrak{x}^2}(I)}{dt} = 2F_{\mathfrak{x}}(\sigma_- + \sigma_+) + F_1(I) = 2F_{\mathfrak{x}}(\sigma_- + \sigma_+) + I, \qquad F_{\mathfrak{x}^2}(I, 0) = 0. \tag{4}$$

Applying Lemma 3.4 to $F_{\mathfrak{x}}(\sigma_- + \sigma_+)$, we obtain

$$\frac{dF_{\mathfrak{x}}(\sigma_- + \sigma_+)}{dt} = -\tfrac{1}{2}F_{\mathfrak{x}}(\sigma_- + \sigma_+) + 2F_1(\sigma_+\sigma_-), \qquad F_{\mathfrak{x}}(\sigma_- + \sigma_+, 0) = 0. \tag{5}$$

Finally, $F_1(\sigma_+\sigma_-)$ satisfies

$$\frac{dF_1(\sigma_+\sigma_-)}{dt} = -F_1(\sigma_+\sigma_-), \qquad F_1(\sigma_+\sigma_-, 0) = \sigma_+\sigma_-. \tag{6}$$

Solving (6), (5) and (4) successively leads first to $F_1(\sigma_+\sigma_-) = e^{-t}\sigma_+\sigma_-$, then to $F_{\mathfrak{x}}(\sigma_- + \sigma_+) = 4(e^{-\frac{t}{2}} - e^{-t})\sigma_+\sigma_-$ and finally to $F_{\mathfrak{x}^2}(I) = 8(e^{-\frac{t}{2}} - 1)^2\sigma_+\sigma_- + tI$. Consequently, the quality of the measurement M of σ_x with pointer $Y = (2 - 2e^{-\frac{t}{2}})^{-1}(A_t + A_t^*)$ is given by

$$\sigma^2 = \|M(Y^2) - M(Y)^2\| = \left\| \frac{8(e^{-\frac{t}{2}} - 1)^2\sigma_+\sigma_- + tI}{(2 - 2e^{-\frac{t}{2}})^2} - I \right\| = \left\| 2\sigma_+\sigma_- + \left(\frac{t}{(2 - 2e^{-\frac{t}{2}})^2} - 1 \right) I \right\|$$
$$= \frac{t}{(2 - 2e^{-\frac{t}{2}})^2} + 1.$$

This expression takes its minimal value 2.228 at $t = 2.513$, leading to a quality $\sigma = 1.493$.

The calculation above has an interesting side product. The observable $M\big((A_t + A_t^*)^2\big)$ depends linearly on σ_z, indicating that in addition to information on σ_x, also information on σ_z in the initial qubit-state ends up in the measurement outcome. Indeed, if we use as a pointer

$$\tilde{Y} := \frac{(A_t + A_t^*)^2 - tI}{4(e^{-\frac{t}{2}} - 1)^2} - I, \tag{7}$$

then we have $M(\tilde{Y}) = \sigma_z$, so that M is also a measurement of σ_z with pointer $\tilde{Y}$.

Note that the pointers Y and $\tilde{Y}$ commute, i.e. measuring $A_t + A_t^*$ via the homodyne detection scheme is an indirect *joint measurement* of σ_x and σ_z. If we would also like to gain some information about σ_y, we could for example sweep the measured quadrature through $[0, 2\pi)$ in time by measuring $e^{i\omega t}A_t + e^{-i\omega t}A_t^*$ instead. In this paper however, we will restrict ourselves to continuous time measurement of $A_t + A_t^*$, as additional information on σ_y would deteriorate the quality of σ_x- and/or σ_z-measurement. The following theorem is a Heisenberg-like relation that gives a bound on how well joint measurements can be performed.

Theorem 3.5: ***(Joint Measurement [18])*** *Let $M: B(\mathcal{F}) \to M_2(\mathbb{C})$ be an unbiased measurement of self-adjoint observables $X \in M_2(\mathbb{C})$ and $\tilde{X} \in M_2(\mathbb{C})$ with self-adjoint commuting pointers Y and $\tilde{Y}$ in (or affiliated to) $B(\mathcal{F})$, respectively. Then for their corresponding qualities σ and $\tilde{\sigma}$ the following relation holds*

$$2\sigma\tilde{\sigma} \geq \|[X, \tilde{X}]\|.$$

Denote by $\tilde{\sigma}$ the quality of the σ_z measurement with the pointer $\tilde{Y}$ defined in (7). Since $[\sigma_x, \sigma_z] = -2i\sigma_y$, the qualities σ and $\tilde{\sigma}$ (corresponding to the pointers Y and $\tilde{Y}$, respectively) satisfy the inequality

$$\sigma\tilde{\sigma} \geq 1. \tag{8}$$

Using similar techniques as before, that is recursively calculating $F_{\mathfrak{X}^4}(I)$ via Lemma 3.4, we find

$$\tilde{\sigma}^2 = \frac{t^2}{8(e^{-\frac{t}{2}} - 1)^4} + \frac{2t - 4(e^{-\frac{t}{2}} - 1)^2}{(e^{-\frac{t}{2}} - 1)^2}.$$

This expression takes its minimal value 8.836 at $t = 2.513$. This leads to a quality $\tilde{\sigma} = 2.973$, which means that $\sigma\tilde{\sigma} = 4.437$, i.e. we are far removed from hitting the bound 1 in (8). However, there is still some room for manoeuvring by post-processing of the homodyne measurement data.

4 The weighted path

Let us presently return to our homodyne detection experiment. We observe $A_\tau + A_\tau^*$ continuously in time, i.e. the result of our measurement is a path ω of measurement results ω_τ (the photocurrent integrated up to time τ) for $A_\tau + A_\tau^*$. This means that we have a space Ω of all possible measurement paths and that we can identify an operator $A_\tau + A_\tau^*$ with the map from Ω to $\mathbb{R}$ mapping a measurement path $\omega \in \Omega$ to the measurement result ω_τ at time τ. That is, we have simultaneously diagonalized the family of commuting operators $\{A_\tau + A_\tau^* |\ \tau \geq 0\}$ and viewed them as random variables on the spectrum Ω. The spectral projectors of the operators $\{A_\tau + A_\tau^* |\ 0 \leq \tau \leq t\}$ endow Ω with a filtration of σ-algebras Σ_t. Furthermore, the states ρ^τ, defined by $\rho^\tau(W) := \rho\otimes\phi(U_\tau^* W U_\tau)$ provide a family of consistent measures $\mathbb{P}_\tau$ on (Ω, Σ_τ), turning it into the probability space $(\Omega, \Sigma_t, \mathbb{P})$. (See e.g. [8].)

We aim to find random variables on $(\Omega, \Sigma_t, \mathbb{P})$ having distributions resembling those of σ_x and σ_z in the initial state ρ. In the previous section we used the random variables

$$Y(\omega) = \frac{\omega_\tau}{2 - 2e^{-\frac{\tau}{2}}} \quad \text{and} \quad \tilde{Y}(\omega) = \frac{\omega_\tau^2 - \tau}{4(e^{-\frac{\tau}{2}} - 1)^2} - 1, \quad \tau = 2.513 \tag{9}$$

for σ_x and σ_z, respectively. Our next goal is to find the optimal random variables, in the sense of the previously defined quality.

4.1 Restricting the class of pointers

In our specific example, M is given by $M(B) = \mathrm{id} \otimes \phi(U_\tau^* I \otimes B U_\tau)$. Note that stochastic integrals with respect to the annihilator A_τ acting on the vacuum vector Φ are zero. Therefore, we can modify U_τ to Z_τ, given by

$$dZ_\tau = \Big\{\sigma_-(dA_\tau^* + dA_\tau) - \tfrac{1}{2}\sigma_+\sigma_- d\tau\Big\} Z_\tau, \quad Z_0 = I,$$

without affecting M [6]. Therefore, for all $B \in B(\mathcal{F})$, we have $M(B) = \mathrm{id} \otimes \phi(U_\tau^* I \otimes B U_\tau) = \mathrm{id} \otimes \phi(Z_\tau^* I \otimes B Z_\tau)$. The solution Z_t can readily be found, it is given by

$$Z_t = \begin{pmatrix} e^{-\frac{1}{2}t} & 0 \\ \int_0^t e^{-\frac{1}{2}\tau}(dA_\tau + dA_\tau^*) & 1 \end{pmatrix}.$$

Note that Z_t, as a matrix valued function of the measurement path, is an element of $M_2(\mathbb{C}) \otimes \mathcal{C}_t$, where $\mathcal{C}_t$ is the commutative von Neumann algebra generated by $A_\tau + A_\tau^*$, $0 \le \tau \le t$. Moreover we see that Z_t is not a function of all the $(A_\tau + A_\tau^*)$'s separately, it is only a function of the endpoint of the weighted path $Y_t = \int_0^t e^{-\frac{1}{2}\tau}(dA_\tau + dA_\tau^*)$ [24]. Therefore if we define $\mathcal{S}_t \subset \mathcal{C}_t$ to be the commutative von Neumann algebra generated by Y_t, then we even have $Z_t \in M_2(\mathbb{C}) \otimes \mathcal{S}_t$.

Denote by $C \mapsto \mathbb{E}[\,C\,|\mathcal{S}_t]$ the unique classical conditional expectation from $\mathcal{C}_t$ onto $\mathcal{S}_t$ that leaves ϕ invariant, i.e. $\phi(\mathbb{E}[\,C\,|\mathcal{S}_t]) = \phi(C)$ for all $C \in \mathcal{C}_t$. We can extend $\mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ by tensoring it with the identity map on the 2×2 matrices to obtain a map $\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ from $M_2(\mathbb{C}) \otimes \mathcal{C}_t$ onto $M_2(\mathbb{C}) \otimes \mathcal{S}_t$. From the positivity of $\mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ as a map between commutative algebras, it follows that $\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ is completely positive. Since $\mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ satisfies $\mathbb{E}[\,CS\,|\mathcal{S}_t] = \mathbb{E}[\,C\,|\mathcal{S}_t]S$ for all $C \in \mathcal{C}_t$ and $S \in \mathcal{S}_t$, we find that $\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ satisfies the *module property*, i.e.

$$\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t](A_1 B A_2) = A_1 \Big(\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t](B)\Big) A_2,$$

for all $A_1, A_2 \in M_2(\mathbb{C}) \otimes \mathcal{S}_t$ and $B \in M_2(\mathbb{C}) \otimes \mathcal{C}_t$. Moreover, if ρ is a state on $M_2(\mathbb{C})$, then it follows from the invariance of ϕ under $\mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ that $\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ leaves $\rho \otimes \phi$ invariant. We conclude that, given ρ on $M_2(\mathbb{C})$, the map $\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ from $M_2(\mathbb{C}) \otimes \mathcal{C}_t$ onto $M_2(\mathbb{C}) \otimes \mathcal{S}_t$ is the unique conditional expectation in the noncommutative sense of [21] that leaves $\rho \otimes \phi$ invariant. We will use the shorthand $\mathbb{E}_{\mathcal{S}_t}$ for $\mathrm{id} \otimes \mathbb{E}[\,\cdot\,|\mathcal{S}_t]$ in the following.

Lemma 4.1: *Let $C \in \mathcal{C}_t$ be a pointer with quality σ_C such that $M(C) = X$. Then $\tilde{C} := \mathbb{E}[\,C\,|\mathcal{S}_t]$ is also a pointer with $M(\tilde{C}) = X$, and with quality $\sigma_{\tilde{C}} \le \sigma_C$.*

Proof. Note that for all states ρ on $M_2(\mathbb{C})$ we have

$$\begin{aligned} \rho\big(M(\tilde{C})\big) &= \rho \otimes \phi\big(Z_t^* I \otimes \tilde{C} Z_t\big) = \rho \otimes \phi\Big(Z_t^* \mathbb{E}_{\mathcal{S}_t}(I \otimes C) Z_t\Big) = \rho \otimes \phi\Big(\mathbb{E}_{\mathcal{S}_t}\big(Z_t^* I \otimes C Z_t\big)\Big) \\ &= \rho \otimes \phi(Z_t^* I \otimes C Z_t) = \rho\big(M(C)\big) = \rho(X), \end{aligned}$$

where we used the module property and the fact that Z_t is an element of $M_2 \otimes \mathcal{S}_t$ in the third step and the invariance of $\rho \otimes \phi$ in the fourth step. Since this holds for all states ρ on $M_2(\mathbb{C})$, we conclude that $M(\tilde{C}) = X$.

As for the variance, we note first that the conditional expectation $\mathbb{E}_{\mathcal{S}_t}$ is a completely positive identity preserving map. Therefore, for all self-adjoint $C \in \mathcal{C}_t$, we have

$$\mathbb{E}_{\mathcal{S}_t}(I \otimes C^2) \ge \Big(\mathbb{E}_{\mathcal{S}_t}(I \otimes C)\Big)^2$$

by the Cauchy-Schwarz inequality for completely positive maps [23].

We can now apply the same strategy as before. For all states ρ on $M_2(\mathbb{C})$ we have

$$\begin{aligned}
\rho\big(M(C^2)\big) &= \rho\otimes\phi(Z_t^* I\otimes C^2 Z_t)\\
&= \rho\otimes\phi\Big(\mathbb{E}_{\mathcal{S}_t}\big(Z_t^* I\otimes C^2 Z_t\big)\Big)\\
&= \rho\otimes\phi\Big(Z_t^*\mathbb{E}_{\mathcal{S}_t}(I\otimes C^2)Z_t\Big)\\
&\geq \rho\otimes\phi\Big(Z_t^*\Big(\mathbb{E}_{\mathcal{S}_t}(I\otimes C)\Big)^2 Z_t\Big)\\
&= \rho\big(M(\tilde{C}^2)\big)\,.
\end{aligned}$$

Thus $M(C^2)\geq M(\tilde{C}^2)$, and in particular $\sigma_C^2 = \|M(C^2)-M(C)^2\| \geq \|M(\tilde{C}^2)-M(\tilde{C})^2\| = \sigma_{\tilde{C}}^2$. $\square$

This has a very useful consequence: if we are looking for pointers that record, say σ_x or σ_z in an optimal fashion, then it suffices to examine only pointers in $\mathcal{S}_t$. Instead of sifting through the collection of all random variables on the measurement outcomes, we are thus allowed to confine the scope of our search to the rather transparent collection of measurable functions of Y_t. In the following, we will look at such pointers $h_t(Y_t)$. We will usually drop the subscript t on h to make the notation lighter.

4.2 Distribution of Y_t

At this point we are interested in the probability distribution of the random variable Y_t. Its characteristic function [4] is given by

$$E(k) := \mathbb{E}_{\rho^t}\big[\exp(-ikY_t)\big] = \rho\otimes\phi\Big(U_t^* I\otimes\exp(-ikY_t)U_t\Big) = \rho\left(F_{\exp(-ik\mathfrak{x})}(I)\right),$$

so that we need only calculate $F_{\exp(-ik\mathfrak{x})}(I)$. For notational convenience we will replace the subscript $\exp(-ik\mathfrak{x})$ by k in the following. Using Lemma 3.4, we find the following system of matrix valued differential equations:

$$\begin{aligned}
\frac{dF_k(I)}{dt} &= -ike^{-\frac{t}{2}}F_k(\sigma_-+\sigma_+) - \frac{k^2e^{-t}}{2}F_k(I),\\
\frac{dF_k(\sigma_-+\sigma_+)}{dt} &= -\tfrac{1}{2}F_k(\sigma_-+\sigma_+) - 2ike^{-\frac{t}{2}}F_k(\sigma_+\sigma_-) - \frac{k^2e^{-t}}{2}F_k(\sigma_-+\sigma_+),\\
\frac{dF_k(\sigma_+\sigma_-)}{dt} &= -F_k(\sigma_+\sigma_-) - \frac{k^2e^{-t}}{2}F_k(\sigma_+\sigma_-),\\
\text{with initially}\quad & F_k(I,0)=I,\quad F_k(\sigma_-+\sigma_+,0)=\sigma_-+\sigma_+,\quad F_k(\sigma_+\sigma_-,0)=\sigma_+\sigma_-.
\end{aligned}$$

Solving this system leads to

$$F_k(I) = e^{-\frac{k^2(1-e^{-t})}{2}}\left(I - ik\big(1-e^{-t}\big)\big(\sigma_-+\sigma_+\big) - k^2\big(1-e^{-t}\big)^2\sigma_+\sigma_-\right).$$

We define the Fourier transform to be $\mathcal{F}(f)(x) := \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} f(k)e^{ikx}dk$. Then the probability density of Y_t with respect to the Lebesgue measure is given by $\frac{1}{\sqrt{2\pi}}\mathcal{F}(E)(x) = \frac{1}{\sqrt{2\pi}}\rho\Big(\mathcal{F}\big(F_k(I)\big)(x)\Big)$.

Defining $p(x) := \frac{1}{\sqrt{2\pi}}\mathcal{F}\big(F_k(I)\big)(x)$, we can write

$$p(x) = \frac{e^{-\frac{1}{2}\frac{x^2}{1-e^{-t}}}}{\sqrt{2\pi(1-e^{-t})}}\Big(I + x(\sigma_- + \sigma_+) + (x^2 - 1 + e^{-t})\sigma_+\sigma_-\Big),$$

i.e. Y_t is distributed according to a Gaussian perturbed by the matrix elements of the initial state $\rho(\sigma_- + \sigma_+) = \rho(\sigma_x)$ and $\rho(\sigma_+\sigma_-) = \frac{1}{2}\rho(\sigma_z) + \frac{1}{2}$. No information about ρ on σ_y enters the distribution though. To gain information about σ_y we would have to change our continuous time measurement setup, as we discussed before. If we absorb a constant $(1-e^{-t})^{-\frac{1}{2}}$ in the definition of Y_t, i.e. $Y_t := (1-e^{-t})^{-\frac{1}{2}}\int_0^t e^{-\frac{\tau}{2}}(dA_\tau + dA_\tau^*)$, then its density becomes

$$p(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}\Big(I + \beta_t y(\sigma_- + \sigma_+) + \beta_t^2(y^2-1)\sigma_+\sigma_-\Big), \tag{10}$$

where $\beta_t := \sqrt{1-e^{-t}}$.

5 Variational calculus

In Lemma 4.1, we have shown that it suffices to consider only random variables of the form $h(Y_t)$ for some measurable h. In equation (10), we have captured the probability distribution of Y_t. All that remains now is to calculate the optimal h, which can be done with variational calculus.

5.1 Optimal σ_x-measurement

We seek the function h^* for which the quality σ of the pointer $h^*(Y_t)$ for σ_x-measurement is optimal. In other words, we need

$$\sigma^2 := \left\| \int_{-\infty}^{\infty} h^2(y)p(y)dy - \Big(\int_{-\infty}^{\infty} h(y)p(y)dy\Big)^2 \right\| := \left\| \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \right\| \tag{11}$$

to be minimal under the restriction $\int_{-\infty}^{\infty} h(y)p(y)dy = \sigma_x$.

Now σ^2 is the norm of a diagonal 2×2-matrix with entries d_1 and d_2. Both depend smoothly on h, but $\sigma^2 = \max\{d_1, d_2\}$ does not. There are three possibilities:

I) $\sigma^2 = d_1$ in some open neighborhood of h^*. To find these h^*, we must minimize the smooth functional d_1 and then check whether $d_1 < d_2$.

II) $\sigma^2 = d_2$ in some open neighborhood of h^*. To find these h^*, we must minimize d_2 and check whether $d_2 < d_1$.

III) $d_1 = d_2$ for h^*. To find these h^*, we must minimize d_1 subject to the condition $d_1 = d_2$.

In principle, we need three different functionals Λ_1, Λ_2 and Λ_3 for these three distinct cases. However, it turns out that we can make due with the following functional

$$\begin{aligned}\Lambda(h,\kappa,\gamma_1,\gamma_2,\gamma_3) := \Big(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h^2(y)e^{-\frac{1}{2}y^2}dy - 1\Big) + \kappa\Big(\frac{\beta_t^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h^2(y)(y^2-1)e^{-\frac{1}{2}y^2}dy\Big) +\\ \gamma_1\Big(\int_{-\infty}^{\infty} h(y)e^{-\frac{1}{2}y^2}dy\Big) + \gamma_2\Big(\int_{-\infty}^{\infty} h(y)(y^2-1)e^{-\frac{1}{2}y^2}dy\Big) +\\ \gamma_3\Big(\frac{\beta_t}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h(y)ye^{-\frac{1}{2}y^2}dy - 1\Big).\end{aligned} \tag{12}$$

The constants γ_1, γ_2 and γ_3 are the Lagrange multipliers enforcing $\int_{-\infty}^{\infty} h(y)p(y)dy = \sigma_x$. These are needed in all cases: Λ_1, Λ_2 and Λ_3. One can readily check that setting $\kappa = 0$ in Λ yields Λ_1, setting $\kappa = 1$ yields Λ_2 and considering κ as a free Lagrange multiplier forces $d_1 = d_2$, so that one has $\Lambda_3 = \Lambda$.

All three cases lead to similar optimality conditions. The requirement that the optimal solution be stable under first order variations yields h^* satisfying either

$$h^*(x) = \frac{C_1x + C_2}{x^2+\varepsilon} + C_3 \tag{13}$$

or

$$h^*(x) = C_4x^2 + C_5x + C_6 \tag{14}$$

for some real constants $C_1, C_2, C_3, C_4, C_5, C_6$ and ε depending on $\kappa, \gamma_1, \gamma_2, \gamma_3$.

Suppose that h^* takes the form (14). The constraint $\int_{-\infty}^{\infty} h^*(y)p(y)dy = \sigma_- + \sigma_+$ will then force $C_4 = C_6 = 0$ and $C_5 = \beta_t^{-1}$, so that $h^*(y) = \beta_t^{-1}y$. The random variable we are investigating is simply the observed path, weighted by the function $f(\tau) = \beta_t^{-1}e^{-\tau/2}$, with t the final time of measurement. Since all the integrals we encounter are Gaussian moments, we can readily compute $M(h^{*2}(Y_t)) = \int_{-\infty}^{\infty} h^{*2}(y)p(y)dy$ to be $2\sigma_+\sigma_- + \beta_t^{-2}I$. Thus

$$\sigma^2 = \|(2\sigma_+\sigma_- + \beta_t^{-2}I) - (\sigma_- + \sigma_+)^2\| = 1 + \beta_t^{-2}\,.$$

For $t \to \infty$, this amounts to $\sigma \to \sqrt{2}$. Already, we have improved on the naive result $\sigma = 1.493$ obtained previously.

We proceed with the more involved case (13), which will provide us with the optimal solution. Before we continue with the constants C_1, C_2, C_3 and ε however, we calculate some integrals for later use.

Definition 5.1: Define the error function $\mathrm{erf}(x)$ and integrals $I(\varepsilon)$ and $J(\varepsilon)$ by

$$\mathrm{erf}(x) := \frac{2}{\sqrt{\pi}}\int_0^x e^{-u^2}du, \quad I(\varepsilon) := \int_{-\infty}^{\infty}\frac{e^{-\frac{x^2}{2}}}{x^2+\varepsilon}dx, \quad J(\varepsilon) := \int_{-\infty}^{\infty}\frac{e^{-\frac{x^2}{2}}}{(x^2+\varepsilon)^2}dx.$$

Lemma 5.2:

$$J(\varepsilon) = \frac{\sqrt{2\pi} + (1-\varepsilon)I(\varepsilon)}{2\varepsilon} \quad \textit{and} \quad I(\varepsilon) = \pi\sqrt{\frac{e^\varepsilon}{\varepsilon}}\left(1 - \mathrm{erf}\left(\sqrt{\frac{\varepsilon}{2}}\right)\right).$$

Proof. Since the Fourier transform of $e^{-\sqrt{\varepsilon}|k|}$ is equal to $\sqrt{\frac{2\varepsilon}{\pi}}\frac{1}{x^2+\varepsilon}$, we find

$$\begin{aligned} I(\varepsilon) &= \sqrt{\frac{\pi}{2\varepsilon}}\int_{-\infty}^{\infty}\mathcal{F}\big(e^{-\sqrt{\varepsilon}|k|}\big)\mathcal{F}\big(e^{-\frac{k^2}{2}}\big)dx = \sqrt{\frac{\pi}{2\varepsilon}}\int_{-\infty}^{\infty}e^{-\sqrt{\varepsilon}|k|}e^{-\frac{k^2}{2}}dk \; = \sqrt{\frac{2\pi}{\varepsilon}}\int_{0}^{\infty}e^{-\sqrt{\varepsilon}k}e^{-\frac{k^2}{2}}dk \\ &= \sqrt{\frac{2\pi}{\varepsilon}}e^{\frac{1}{2}\varepsilon}\int_{\sqrt{\varepsilon}}^{\infty}e^{-\frac{u^2}{2}}du = \pi\sqrt{\frac{e^{\varepsilon}}{\varepsilon}}\left(1-\operatorname{erf}\left(\sqrt{\frac{\varepsilon}{2}}\right)\right), \end{aligned}$$

where, in the second step, we have used that the Fourier transform $\mathcal{F}$ is unitary. The expression for J follows from

$$\begin{aligned} 0 = \left.\frac{xe^{-\frac{x^2}{2}}}{x^2+\varepsilon}\right|_{-\infty}^{\infty} &= \int_{-\infty}^{\infty}\frac{d}{dx}\left(\frac{xe^{-\frac{x^2}{2}}}{x^2+\varepsilon}\right)dx = \int_{-\infty}^{\infty}\left(\frac{1-x^2}{x^2+\varepsilon}-\frac{2x^2}{(x^2+\varepsilon)^2}\right)e^{-\frac{x^2}{2}}dx \\ &= \int_{-\infty}^{\infty}\left(-1+\frac{\varepsilon-1}{x^2+\varepsilon}+\frac{2\varepsilon}{(x^2+\varepsilon)^2}\right)e^{-\frac{x^2}{2}}dx = -\sqrt{2\pi}+(\varepsilon-1)I(\varepsilon)+2\varepsilon J(\varepsilon). \end{aligned}$$

□

The condition $\int_{-\infty}^{\infty}h^*(y)p(y)dy = \sigma_- + \sigma_+ = \sigma_x$ implies

$$C_1 = \frac{\sqrt{2\pi}}{\beta_t\big(\sqrt{2\pi}-\varepsilon I(\varepsilon)\big)}, \qquad C_2 = C_3 = 0$$

which fixes C_1 as a function of ε. The next step is to express d_1 and d_2 in terms of ε:

$$\begin{aligned} d_2 &= \frac{C_1^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{y^2}{(y^2+\varepsilon)^2}e^{-\frac{y^2}{2}}dy - 1 = \frac{C_1^2}{\sqrt{2\pi}}\Big(I(\varepsilon)-\varepsilon J(\varepsilon)\Big) - 1, \\ d_1 &= \frac{C_1^2\beta_t^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty}\frac{y^2(y^2-1)}{(y^2+\varepsilon)^2}e^{-\frac{y^2}{2}}dy + d_2 = \frac{C_1^2\beta_t^2}{\sqrt{2\pi}}\Big(\sqrt{2\pi}-(1+2\varepsilon)I(\varepsilon)+\varepsilon(1+\varepsilon)J(\varepsilon)\Big) + d_2. \end{aligned}$$

First, we use Lemma 5.2 to express the above in terms of elementary functions and the error function. Then, using Maple, we find that $\varepsilon \mapsto \max\{d_1, d_2\}$ has a unique minimum at $\varepsilon = 0.605$, for which $d_1 = d_2 = 0.470$. This leads to a C_1 that equals 2.359, and to a quality of

$$\sigma = \sqrt{\max\{d_1, d_2\}} = 0.685\,.$$

5.2 Optimal σ_z-measurement

For optimal σ_z-measurement, we can run the same program. We search for the function $\tilde{h}$ which optimizes the quality $\tilde{\sigma}$, under the restriction that $\tilde{h}(Y_t)$ be a pointer for σ_z-measurement. That is, we search for a function $\tilde{h}$ minimizing the functional of equation (11), but now under the restriction $\int_{\infty}^{\infty} h(y)p(y)dy = \sigma_z$. Again there are three cases of interest, $d_1 = d_2$, $d_1 > d_2$ and $d_2 > d_1$, which

we can treat simultaneously by introducing, analogous to equation (12), the functional

$$\begin{aligned}\tilde{\Lambda}(h,\kappa,\gamma_1,\gamma_2,\gamma_3) := &\Big(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h^2(y)e^{-\frac{1}{2}y^2}dy - 1\Big) + \kappa\Big(\frac{\beta_t^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h^2(y)(y^2-1)e^{-\frac{1}{2}y^2}dy\Big) + \\ &\gamma_1\Big(\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h(y)e^{-\frac{1}{2}y^2}dy + 1\Big) + \gamma_2\Big(\frac{\beta_t^2}{\sqrt{2\pi}}\int_{-\infty}^{\infty} h(y)(y^2-1)e^{-\frac{1}{2}y^2}dy - 2\Big) + \\ &\gamma_3\Big(\int_{-\infty}^{\infty} h(y)ye^{-\frac{1}{2}y^2}dy\Big).\end{aligned}$$

Indeed, γ_1, γ_2 and γ_3 are the Lagrange multipliers enforcing the restriction $\int_{\infty}^{\infty} h(y)p(y)dy = \sigma_z$. Again, the functional σ^2 of equation (11) depends non-differentiably on h when $d_1 = d_2$. We then have to search for the optimum among the points of non-differentiability, in which case κ is the Lagrange multiplier confining us to these points. If $d_1 > d_2$ then $\kappa = 1$ and if $d_2 > d_1$ then $\kappa = 0$. Summarizing, wherever Λ takes its minimal value, optimality implies $\frac{\delta\tilde{\Lambda}}{\delta h}(\tilde{h}, \kappa, \gamma_1, \gamma_2, \gamma_3) = 0$ for some $\kappa, \gamma_1, \gamma_2$ and γ_3. Performing the functional derivative yields either

$$\tilde{h}(x) = \frac{D_1 x + D_2}{x^2 + \delta} + D_3 \tag{15}$$

or

$$\tilde{h}(x) = D_4 x^2 + D_5 x + D_6 \tag{16}$$

for some (time-dependent) constants $D_1, D_2, D_3, D_4, D_5, D_6$ and δ depending on $\kappa, \gamma_1, \gamma_2$ and γ_3.

Again, we begin with the least demanding case (16), resulting from $\kappa = 0$. The condition $\int_{-\infty}^{\infty} \tilde{h}(y)p(y)dy = \sigma_z$ implies $D_5 = 0$, $D_4 = \beta_t^{-2}$ and $D_6 = -1 - \beta_t^2$. For $t \to \infty$, this leads to

$$\sigma^2 = \|M(\tilde{h}^2(Y_t)) - M(\tilde{h}(Y_t))^2\| = \|(4\sigma_+\sigma_- + 3I) - I\| = 6\,,$$

so that $\sigma \to \sqrt{6}$.

This improves the result $\tilde{\sigma} = 2.973$ obtained previously, but once again the ultimate bound will be reached in the more arduous case (15). There, the condition $\int_{-\infty}^{\infty} \tilde{h}(y)p(y)dy = \sigma_z$ implies

$$D_1\big(\sqrt{2\pi} - \delta I(\delta)\big) = 0, \qquad D_2 = \frac{2\sqrt{2\pi}}{\beta_t^2\big(\sqrt{2\pi} - (1+\delta)I(\delta)\big)}, \qquad D_3 = -\frac{\sqrt{2\pi} + I(\delta)D_2}{\sqrt{2\pi}}.$$

This leads to expressions for d_1 and d_2 as a function of δ. Using Lemma 5.2 and Maple once more, we find that the function $\delta \mapsto \max\{d_1, d_2\}$ has a unique minimum at $\delta = 2.701$, for which $d_1 = d_2 = 2.373$. This leads to a quality of

$$\tilde{\sigma} = \sqrt{\max\{d_1, d_2\}} = 1.540,$$

attained for $D_1 = 0$, $D_2 = -21.649$ and $D_3 = 5.391$. For the joint measurement this leads to

$$\sigma\tilde{\sigma} = 1.056\,.$$

Although we did not achieve the bound of 1 provided by Theorem 3.5, we have come as close as the measurement setup allows. We conclude that, using the setup investigated in this article, no simultaneous measurement of σ_x and σ_z will be able to approach the quantum bound by more than 5.6 %. Furthermore, we have identified the unique pointers for this optimal measurement in equations (13) and (15).

6 Distribution of pointer variables

We have designed pointers $h^*(Y_t)$ and $\tilde{h}(Y_t)$ in such a way that their distributions in the final state best resemble the distributions of σ_x and σ_z in the initial state. We will now calculate and plot these final densities.

6.1 Calculation of h^*- and $\tilde{h}$-densities

Let ρ be the initial state of the qubit and let it be parameterized by its Bloch vector (P_x, P_y, P_z). By equation (10), the density $q(y)$ of Y_t is given by

$$q(y) = \rho\big(p(y)\big) = \frac{e^{-\frac{1}{2}y^2}}{\sqrt{2\pi}}\Big(1 + \beta_t y P_x + \beta_t^2 (y^2 - 1)\frac{P_z + 1}{2}\Big). \tag{17}$$

We are interested in the the distributions $r(x)$ and $s(x)$ of $h^*(Y_t)$ and $\tilde{h}(Y_t)$ respectively. Let us start with h^*. From equation (13), we first calculate the points y where $h^*(y) = x$ for some fixed value of x.

$$y_\pm = \frac{C_1 \pm \sqrt{C_1^2 - 4x^2\varepsilon}}{2x}$$

By the Frobenius-Peron equation (see e.g. [19]), $r(x)$ is given by

$$r(x) = \sum_{+,-} \frac{q(y_\pm)}{|{h^*}'(y_\pm)|},$$

which leads immediately to

$$r(x) = \sum_{+,-} \frac{(y_\pm^2 + \varepsilon)^2\Big(1 + \beta_t y_\pm P_x + \beta_t^2 (y_\pm^2 - 1)\frac{P_z+1}{2}\Big)}{C_1|y_\pm^2 - \varepsilon|}\frac{e^{-\frac{1}{2}y_\pm^2}}{\sqrt{2\pi}}, \tag{18}$$

where it is understood that $r(x) \neq 0$ only for $x \in [-\frac{C_1}{2\sqrt{\varepsilon}}, \frac{C_1}{2\sqrt{\varepsilon}}]$.

We run a similar analysis for $s(x)$. The points y in which $\tilde{h}(y) = x$ are given by

$$y_\pm = \pm\sqrt{\frac{(x - D_3)\delta - D_2}{D_3 - x}}.$$

This leads to

$$s(x) = \sum_{+,-} \frac{(y_\pm^2 + \delta)^2\Big(1 + \beta_t y_\pm P_x + \beta_t^2 (y_\pm^2 - 1)\frac{P_z+1}{2}\Big)}{2|D_2 y_\pm|}\frac{e^{-\frac{1}{2}y_\pm^2}}{\sqrt{2\pi}}, \tag{19}$$

with $s(x) \neq 0$ only for $x \in [D_3 + \frac{D_2}{\delta}, D_3]$. We proceed with a graphical illustration of the results obtained so far.

6.2 Plots of σ_x-measurement

According to formula 17, the distribution of the endpoint of the weighted path depends on the input qubit-state. For instance, the negative σ_x-eigenstate, the tracial state and the positive σ_x-eigenstate lead to the distributions below:

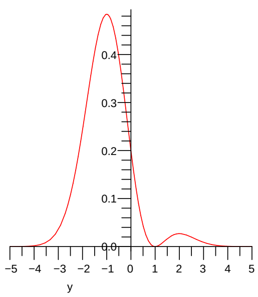

Figure 1: Probability density of the endpoint of the weighted path for input $|\leftarrow\rangle$.

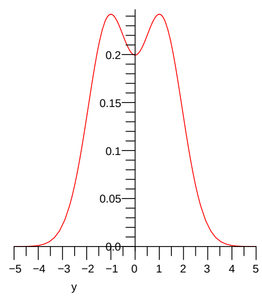

Figure 2: Probability density of the endpoint of the weighted path for input tr.

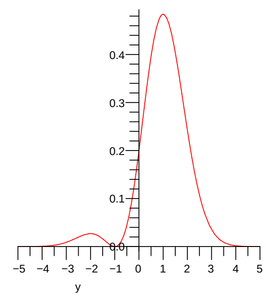

Figure 3: Probability density of the endpoint of the weighted path for input $|\rightarrow\rangle$.

In order to estimate σ_x, we use the pointer of σ_x given by

$$h^*(x) = \frac{C_1 x}{x^2 + \varepsilon}, \tag{20}$$

with $\varepsilon = 0.605$ and $C_1 = 2.359$. It is illustrated in figure 4 to the right.
In formula (18), we have calculated the probability distributions of this pointer under the distributions of the endpoint of the weighted path illustrated in figures 1, 2 and 3. They are illustrated in figures 5, 6 and 7 below.

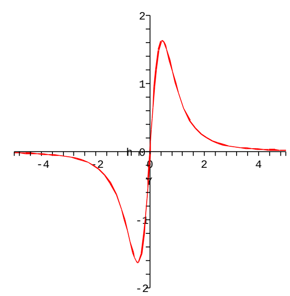

Figure 4: Pointer for σ_x

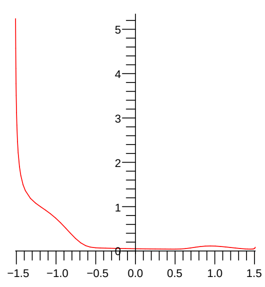

Figure 5: Probability density of the σ_x-pointer for input $|\leftarrow\rangle$.

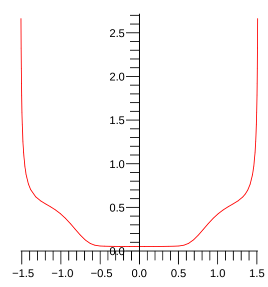

Figure 6: Probability density of the σ_x-pointer for input tr.

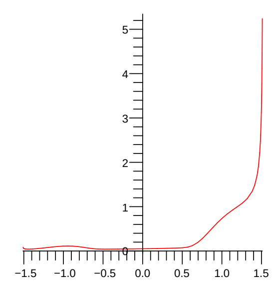

Figure 7: Probability density of the σ_x-pointer for input $|\rightarrow\rangle$.

6.3 Plots of σ_z-measurement

We repeat this for the σ_z-pointer. By formula 17, the positive σ_z-eigenstate, the tracial state and the negative σ_z-eigenstate lead to the distributions of the endpoint of the weighted path shown below:

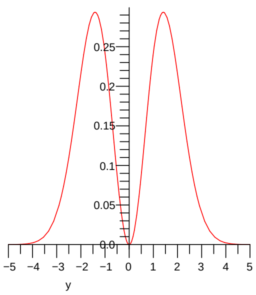

Figure 8: Probability density of the endpoint of the weighted path for input $|\uparrow\rangle$.

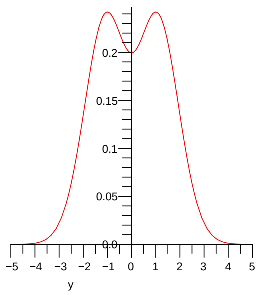

Figure 9: Probability density of the endpoint of the weighted path for input tr.

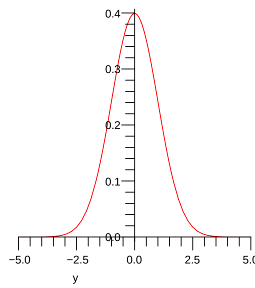

Figure 10: Probability density of the endpoint of the weighted path for input $|\downarrow\rangle$.

In order to estimate σ_z, we use the pointer of σ_z illustrated here to the right. It is given by

$$\tilde{h}(x) = \frac{D_2}{x^2 + \delta} + D_3, \tag{21}$$

with $\delta = 2.701$, $D_2 = -21.649$ and $D_3 = 5.391$.

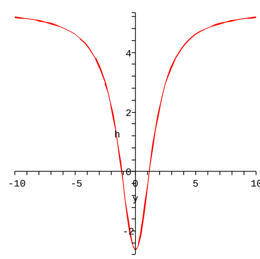

Figure 11: Pointer for σ_z.

From formula (19), we read off the probability distributions of this pointer under the distributions of the endpoint of the weighted path illustrated in figures 8, 9 and 10. They are shown below:

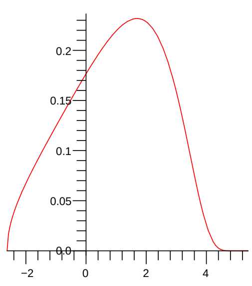

Figure 12: Probability density of the σ_z-pointer for input $|\uparrow\rangle$.

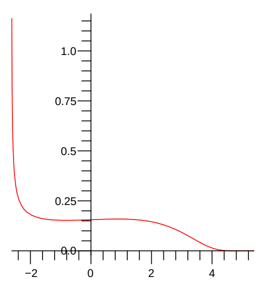

Figure 13: Probability density of the σ_z-pointer for input tr.

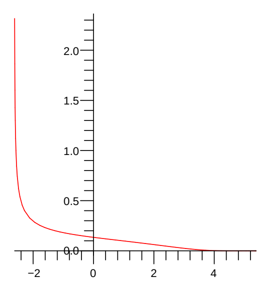

Figure 14: Probability density of the σ_z-pointer for input $|\downarrow\rangle$.

7 Discussion

In this paper, we have investigated homodyne detection of spontaneous decay of a two-level atom into the electromagnetic field. We have seen how the photocurrent, besides carrying information on σ_x (which is immediate from the innovations term in the filtering equation), also carries information on σ_z. Homodyne detection can thus be viewed as a joint measurement of the non-commuting observables σ_x and σ_z in the initial state of the qubit, and we have identified the optimal pointers for this procedure in equations (20) and (21).

One particular feature of the pointers we constructed might seem counterintuitive at first: they yield values outside $[-1, 1]$ with nonzero probability. This is a direct result of our requirement that the measurement be unbiased. Suppose, for example, that the input state is $|\uparrow\rangle$, so that σ_z has value 1. Since the photocurrent carries information on σ_x as well, its information on σ_z is certainly flawed, and will yield estimates $\sigma_z < 1$ at least some of the time. Unbiasedness then implies that also estimates $\sigma_z > 1$ must occur.

On the other hand, an unbiased measurement will yield *on average* the 'true' value of σ_z for any possible input state. (Not just for the 3 possibilities sketched on page 17.) In repeated experiments, optimality of our pointers guarantees fast convergence to these averages.

Theorem 3.5 provides a theoretical bound for the quality of joint measurement of σ_x and σ_z. No conceivable measurement procedure can ever achieve $\sigma\tilde{\sigma} < 1$. It is now clear that this bound cannot be met by way of homodyne detection: a small part of the information extracted from the atom is simply lost in this particular procedure. Constructing the optimal pointers on the photocurrent does yield $\sigma\tilde{\sigma} = 1.056$ however, a figure much closer to the bound than the 4.437 provided by the naïve choice of (9).

Acknowledgment

We thank Mădălin Guţă and Hans Maassen for interesting discussions. We thank John Gough for a critical reading of the text. L.B. acknowledges support by the ARO under Grant DAAD19-03-1-0073.

References

[1] L. Accardi, A. Frigerio, and Y. Lu. The weak coupling limit as a quantum functional central limit. *Commun. Math. Phys.*, 131:537–570, 1990.

[2] M. Armen, J. Au, J. Stockton, A. Doherty, and H. Mabuchi. Adaptive homodyne measurement of optical phase. *Phys. Rev. Lett.*, 89:133602, 2002.

[3] A. Barchielli. Direct and heterodyne detection and other applications of quantum stochastic calculus to quantum optics. *Quantum Opt.*, 2:423–441, 1990.

[4] A. Barchielli and G. Lupieri. Quantum stochastic calculus, operation valued stochastic processes, and continual measurements in quantum mechanics. *J. Math. Phys.*, 26:2222–2230, 1985.

[5] V. Belavkin. Nondemolition stochastic calculus in Fock space and nonlinear filtering and control in quantum systems. In R. Guelerak and W. Karwowski, editors, *Proceedings XXIV Karpacz winter school*, Stochastic methods in mathematics and physics, pages 310–324. World Scientific, Singapore, 1988.

[6] V. Belavkin. Quantum continual measurements and a posteriori collapse on CCR. *Commun. Math. Phys.*, 146:611–635, 1992.

[7] V. Belavkin. Quantum stochastic calculus and quantum nonlinear filtering. *Journal of Multivariate Analysis*, 42:171–201, 1992.

[8] L. Bouten. *Filtering and Control in Quantum Optics*. PhD thesis, Radboud Universiteit Nijmegen, arXiv:quant-ph/0410080, 2004.

[9] L. Bouten, M. Guţă, and H. Maassen. Stochastic Schrödinger equations. *J. Phys. A*, 37:3189–3209, 2004.

[10] H. Carmichael. *An Open Systems Approach to Quantum Optics*. Springer-Verlag, Berlin Heidelberg New-York, 1993.

[11] E. Davies. Quantum stochastic processes. *Commun. Math. Phys.*, 15:277–304, 1969.

[12] E. Davies. *Quantum Theory of Open Systems*. Academic Press, London New-York San Francisco, 1976.

[13] J. Geremia, J. Stockton, and H. Mabuchi. Real-time quantum feedback control of atomic spin-squeezing. *Science*, 304:270–273, 2004.

[14] J. Gough. Feynman diagrams and the quantum stochastic calculus. arXiv:quant-ph/0411064, Nottingham Trent University, 2004.

[15] J. Gough. Quantum flows as Markovian limit of emission, absorption and scattering interactions. *Commun. Math. Phys.*, 254:489–512, 2005.

[16] A. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland Publishing Company, Amsterdam-New York-Oxford, 1982.

[17] R. Hudson and K. Parthasarathy. Quantum Itô's formula and stochastic evolutions. *Commun. Math. Phys.*, 93:301–323, 1984.

[18] B. Janssens. Quantum Measurement. A Coherent Description. arXiv:quant-ph/0503009, Master's Thesis, Radboud Universiteit Nijmegen, 2005.

[19] E. Ott. *Chaos in Dynamical Systems*. Cambridge University Press, 1993.

[20] K. Parthasarathy. *An Introduction to Quantum Stochastic Calculus*. Birkhäuser, Basel, 1992.

[21] M. Takesaki. Conditional expectations in von Neumann algebras. *J. Funct. Anal.*, 9:306–321, 1971.

[22] J. von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin, 1932.

[23] R. F. Werner. Quantum information theory - an invitation. *Springer Tracts in Modern Physics*, 173:14–57, 2001.

[24] H. Wiseman. Quantum trajectories and quantum measurement theory. *Quantum Semiclass. Opt.*, 8:205–222, 1996.

[25] H. Wiseman and G. Milburn. Quantum theory of field-quadrature measurements. *Phys. Rev. A*, 47:642–662, 1993.